An Intraoperative Glucose Control Benchmark for Formal Verification

Sanjian Chen [*], **Matthew O'Kelly** [**], **James Weimer** [*], **Oleg Sokolsky** [*], **Insup Lee** [*]

[*] *Dept. of Computer and Information Science, University of Pennsylvania, Philadelphia, USA (e-mail: sanjian—weimerj—sokolsky—lee@seas.upenn.edu).*
[**] *Dept. of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, USA (e-mail:mokelly@seas.upenn.edu).*

Abstract: Diabetes associated complications are affecting an increasingly large population of hospitalized patients. Since glucose physiology is significantly impacted by patient-specific parameters, it is critical to verify that a clinical glucose control protocol is safe across a wide patient population. A safe protocol should not drive the glucose level into dangerous low (hypoglycemia) or high (hyperglycemia) ranges. Verification of glucose controllers is challenging due to the high-dimensional, non-linear glucose physiological models which contain both unobservable states and unmeasurable patient-specific parameters. This paper presents a hybrid system model of a closed-loop physiological system that includes an existing FDA-accepted high-fidelity physiological model tailored to intraoperative settings and a validated improvement to a clinical glucose control protocol for diabetic cardiac surgery patients. We propose the closed-loop model as a physiological system benchmark for verification and present our initial results on verifying the system using the SMT-based hybrid system verification tool dReach.

Keywords: Formal verification; Medical applications; Safety analysis; Closed-loop Controllers; Glucose control;

1. INTRODUCTION

For the more than 29 million Americans who have diabetes, the risk of death is nearly twice as high when compared to age-matched non-diabetic individuals (Xu et al. (2010)). Those suffering from this disease, especially Type 1 diabetics, depend on insulin self-injections to manage their blood glucose level. As such, glucose regulation is a safety-critical control task: too much insulin causes life-threatening hypoglycemia (low glucose levels) and too little insulin causes hyperglycemia (high glucose levels), a condition that has potential outcomes such as blindness and nerve damage.

While outpatient glucose management has been the primary focus of recent diabetes research (e.g., the artificial pancreas (AP) Cobelli et al. (2011)), mounting evidence suggests that diabetes associated complications among hospitalized patients are increasing (Wallymahmed et al. (2005)); thus, methods for inpatient glycemic control are important (Bruno et al. (2008); McAlister et al. (2005)). During surgeries, patients can suffer from stress-induced glucose fluctuations (Bochicchio et al. (2005)). Data suggests that specialized inpatient glucose level management within a safe range can minimize the hypoglycemia risk and improve clinical outcomes (Subramaniam et al. (2009); Lazar et al. (2004)). Clinicians currently follow rule-based protocols to administer insulin and glucose during surgeries (e.g., see Kohl et al. (2013)), but those protocols are still far from foolproof (Meijering et al. (2006)). Thus, verifying that intraoperative glycemic controllers avoid severe hypoglycemia/hyperglycemia events across a diabetic population is imperative.

Recently, the United States (US) Food and Drug Administration (FDA) has accepted the UVa/Padova Type 1 Diabetes Mellitus Metabolic Simulator (T1DMS) as a substitute for animal testing in certain pre-clinical trials of glucose controllers (Kovatchev et al. (2009); Dalla Man et al. (2014)). The T1DMS utilizes a high-dimensional, multi-modal, and non-linear model with over 30 patient-dependent parameters that are (mostly) unobservable in ordinary T1D patients through standard medical tests. Existing work on evaluating controllers using T1DMS relies on simulating the physiological models with a finite set (typically 300, see Kovatchev et al. (2009)) of "virtual subjects", which are discrete realizations of the model parameters identified through invasive experiments (Basu et al. (2003)). However, there is no formal guarantee that the "virtual subject" set covers the entire T1D population. To this end, formal verification of controllers can provide a new level of safety assurance to clinical practitioners before performing human clinical trials. [1]

This paper makes the following contributions towards formal verification of intraoperative glycemic control. First,

[1] Currently, model-based trials are only approved to replace pre-clinical testing. It is unclear whether model-based trials will ever be approved to replace clinical (human) testing due to unmodeled physiology and comorbidity inherent in all models.

we introduce the model of the closed-loop intraoperative glycemic control system as a case study verification benchmark: the model contains both an FDA-accepted high-fidelity physiological model and a validated intraoperative glycemic control protocol. We also provide over-approximated value ranges of all model states and parameters whose ranges are supported by extensive clinical studies. Second, we implement the benchmark in a recently proposed SMT-based hybrid system verification tool, dReal/dReach (Gao et al. (2013a)). Third, we present a proof-of-concept safety verification of the intraoperative glycemic control benchmark over a non-scalar subspace of each physiological parameter/state.

The rest of this paper is organized as follows: Section 2 presents the problem formulation; Section 3 introduces the diabetes model in the intraoperative setting; Section 4 presents the hybrid system model of the closed-loop physiological system; Section 5 describes the case study of verifying an intra-operative glycemic controller on the surgical physiological model using dReach and includes a presentation of our initial verification results in a subspace of the entire parameter and initial condition range; Section 6 discusses our future work.

2. PROBLEM FORMULATION

In this section, we define the safety verification problem considered in this work. We represent the combined intraoperative glucose control protocol and physiological dynamics (defined in Section 3) as a standard hybrid system,

$$\mathcal{H} = \langle \mathcal{X}, \mathcal{Q}, \mathcal{X}_{init}, \mathcal{X}_{inv}, \mathcal{F}(\mathcal{P}), T \rangle , \tag{1}$$

where $\mathcal{X}$ represents the continuous states, $\mathcal{Q}$ denotes the discrete modes, $\mathcal{X}_{init} \in \mathcal{R}_{\mathcal{X}}$ specifies the initial condition space, $\mathcal{F}(\mathcal{P})$ captures the flows parameterized by a vector $\mathcal{P} \in \mathcal{R}_P$, $\mathcal{X}_{inv}$ identifies invariants mapping modes to flows, and T relates the transitions between modes. A measurable output $y = \phi(t; \mathcal{X}_{init})$ denotes the glucose value, with $\phi(t, \mathcal{X}_{init})$ describing the measurement at time $t \in [0, t_{\max}]$ [2], having evolved from initial condition $\mathcal{X}_{init}$. In this paper, we aim to solve the following safety verification problem:

$$\forall t \in [0, t_{\max}] , \ \forall \mathcal{P} \in \mathcal{R}_P, \ \forall \mathcal{X}_{init} \in \mathcal{R}_{\mathcal{X}}, \quad y \notin \mathcal{R}_{unsafe},$$

where $\mathcal{R}_{unsafe}$ is a region representing unsafe blood glucose content (i.e., hypoglycemia and hyperglycemia).

3. MODELING OF SURGICAL GLUCOSE CONTROL

In this section, we introduce the FDA-accepted T1DMS model (Man et al. (2007); Dalla Man et al. (2014)) modified for the intraoperative clinical scenario and a clinically validated glucose control protocol (Kohl et al. (2013)).

3.1 Glucose-Insulin System Model

The full T1DMS model contains three sub-models (insulin, glucose, and carbohydrate-ingestion) with 13 states and 32 parameters. The original publications (e.g., Man et al. (2007); Kovatchev et al. (2009)) discuss the details of physiological modeling and our previous paper (Chen et al. (2015)) summarizes the model equations from the literature. Since intraoperative patients receive insulin and glucose via intravenous infusion, the two subcutaneous insulin compartment states and the entire carbohydrate-ingestion sub-system can be neglected, resulting in a 7-state intraoperative model, as described in the remainder of this subsection.

The intraoperative model contains an insulin sub-model and a glucose sub-model. The insulin system is a 5-state linear model driven by the insulin input, $u(t)$, written as

$$\dot{I}_p(t) = -(m_2 + m_4) I_p(t) + m_1 I_l(t) + u(t) * 10^2 / BW \tag{2a}$$
$$\dot{X}(t) = P_{2U}/V_i I_p(t) - P_{2U} X(t) - P_{2U} * I_b \tag{2b}$$
$$\dot{I}_1(t) = k_i/V_i I_p(t) - k_i I_1(t) \tag{2c}$$
$$\dot{I}_d(t) = k_i I_1(t) - k_i I_d(t) \tag{2d}$$
$$\dot{I}_l(t) = m_2 * I_p(t) - (m_1 + m_3) I_l(t). \tag{2e}$$

The $I_p(t)$ and $I_l(t)$ states represent insulin mass in the plasma and liver, respectively. $I_1(t)$ and $I_d(t)$ represent a delayed insulin transportation process. $X(t)$ represents an insulin signal in the remote tissue that governs glucose concentration in the interstitial compartment. The model contains a set of parameters that are patient dependent: $m_{1...4}$ and P_{2u} are rates of insulin mass diffusion among different compartments, V_i is the insulin distribution volume, and BW is the body weight.

The glucose system has two states and is written as

$$\begin{aligned} \dot{G}_p(t) = &- k_1 * G_p(t) + k_2 * G_t(t) - F_{snc} + m(t) * 10^3 / BW \\ &+ \max(0, k_{p1} - k_{p2} * G_p(t) - k_{p3} * I_d(t)) \\ &- 1 - \max(0, k_{e1} * (G_p(t) - k_{e2})) \end{aligned} \tag{3a}$$

$$\dot{G}_t(t) = - \frac{(V_{m0} + V_{mx} * X(t)) * G_t(t)}{K_{m0} + G_t(t)} + k_1 * G_p(t) - k_2 * G_t(t) \tag{3b}$$

where, $G_p(t)$ and $G_t(t)$ represent the glucose concentration in plasma and interstitial fluids, respectively. The $G_p(t)$ derivative (Equation 3a) contains two saturation switches $\max(0, k_{p1} - k_{p2} * G_p(t) - k_{p3} * I_d(t))$ and $\max(0, k_{e1} * (G_p - k_{e2}))$, which represent the endogenous glucose production (EGP) and renal glucose clearance, respectively. These two max switches yield four discrete modes in the hybrid system representation of the model, and transitions among the four modes are governed by saturations of the two max terms. The G_t derivative contains a non-linear term $-\frac{(V_{m0} + V_{mx} * X(t)) * G_t(t)}{K_{m0} + G_t(t)}$ that represents the remote insulin signal $X(t)$'s impact on glucose dynamics. The model contains two population static parameters k_{e1} (glomerular filtration rate) and k_{e2} (renal threshold of glucose). All other parameters are patient dependent: k_1 and k_2 are the mass exchange rate between the G_p and G_t compartments; k_{p1} is the extrapolated EGP; k_{p2} is the liver glucose effectiveness; k_{p3} is the insulin action on liver; V_{m0}, V_{mx}, and K_{m0} are model parameters that govern the insulin action on G_t; V_g is the glucose distribution volume. $m(t)$ is the intravenous glucose input into the plasma compartment.

The 7-state intraoperative glucose control model is observed through $y(t) = G_p(t)/V_g$, corresponding to the plasma glucose measurement (in mg/dL). Most of the patient-dependent parameters, except for a few such as the body weight, are not measurable in standard hospital tests. Estimating those parameters on individual patients involves invasive and costly procedures such as the triple-

[2] $t_{\max}$ represents the maximum time the patient is in surgery.

Table 1. Over-Approximated Ranges of the T1DMS Model States

States	Ranges	Units	Example Nominal Value
I_p	$[0, 30]$	pmol/kg	5
X'	$[-500, 500]$	pmol/liter	30
I_1	$[0, 300]$	pmol/liter	120
I_d	$[0, 300]$	pmol/liter	120
I_l	$[0, 30]$	pmol/kg	3
G_p	$[0, 1000]$	mg/kg	200
G_t	$[0, 1000]$	mg/kg	150

Table 2. Over-Approximated Ranges of the T1DMS Model Parameters

Parameters	Ranges	Units	Example Nominal Value
m_1	$[0.1, 1]$	min^{-1}	0.2
m_2	$[0.1, 1]$	min^{-1}	0.3
m_3	$[0.1, 1]$	min^{-1}	0.3
m_4	$[0.05, 0.5]$	min^{-1}	0.1
k_i	$[0.001, 0.02]$	min^{-1}	0.01
P_{2u}	$[0.01, 0.1]$	min^{-1}	0.03
V_i	$[0.02, 0.1]$	liter/kg	0.06
I_b	$[0, 300]$	pmol/liter	100
BW	$[0, 300]$	kg	90
k_1	$[0.02, 0.1]$	min^{-1}	0.05
k_2	$[0.05, 0.3]$	min^{-1}	0.1
k_{p1}	$[1, 10]$	mg/kg/min	5
k_{p2}	$[0.0001, 0.01]$	min^{-1}	0.004
k_{p3}	$[0.001, 0.03]$	mg/kg/min per pmol/liter	0.01
V_{m0}	$[1, 10]$	mg/kg/min	5
V_{mx}	$[0.01, 0.15]$	mg/kg/min per pmol/liter	0.05
K_{m0}	$[100, 1000]$	mg/kg	200
V_g	$[1, 5]$	dL/kg	2

tracer meal protocol experiment (Basu et al. (2003); Man et al. (2007)), which is clearly not feasible in surgical settings. The FDA-accepted T1DMS simulator comes with 10 adult virtual subjects, each of which is a whole realization of the parameters. Those virtual subjects are extracted from the same distribution as the 100 FDA-accepted adult virtual subjects for black-box controller evaluation were.

All the states and parameters in the FDA-accepted model have physiological meanings, and numerous clinical studies have investigated the ranges of values across different populations (Harris et al. (1987); Danaei et al. (2011); Kulcu et al. (2003); Katz et al. (2000); Laakso (1993)). Table 1 lists over-approximated ranges and the units of the seven states and Table 2 lists over-approximated ranges of the eighteen parameters.

3.2 A Proportional-Derivative Glucose Control Protocol

In surgery rooms, clinicians periodically sample the glucose values approximately every 30 minutes, and adjust insulin or glucose inputs only at sample times based on rules defined in the clinical protocols. The insulin input has two types: the continuous intravenous infusion rate $uc(k)$, which will remain constant within a sample period, and insulin bolus $ub(k)$ that is an impulse input. The insulin input $u(t)$ that goes into the plasma insulin compartment $I_p(t)$ in Equation 2a is $(uc(t) + ub(t))$. The glucose input $m(t)$ is in the form of dextrose bolus that is an impulse input to the plasma glucose compartment $G_p(t)$ in Equation 3a.

Our team has collaborated with clinicians in the Division of Critical Care at the Hospital of the University of Pennsylvania in order to evaluate an intraoperative insulin protocol (IIP) that manages the glucose level for cardiac surgery patients (Kohl et al. (2013)). Our previous work identifies the weaknesses of the IIP and proposes a proportional-derivative (PD) controller that reduces intraoperative hypoglycemia while preserving the IIP's strengths in simulation studies. The clinicians who developed the IIP believed that the results warrant prospective in-vivo evaluations of the PD controller (Kohl et al. (2013)).

In this paper, we present a proof-of-concept safety verification of the PD controller for a non-scalar subspace of each physiological parameter/state. The PD controller (see Kohl et al. (2013)) calculates the insulin or glucose dose based on two plasma glucose readings: the current value $y(k)$ and the last reading $y(k-1)$ sampled 30 minutes before; those are the same glucose inputs required by the IIP. The PD controller updates $uc(k)$, $ub(k)$, and $m(k)$ based on $y(k)$ and $y(k-1)$ according to the rules defined in Table 3. The controller gains are static and tuned to minimize the hypoglycemia risk while maximizing quality of glucose control in a T1DMS simulation study. [3]

4. A HYBRID SYSTEM MODEL OF THE PHYSIOLOGY AND CONTROLLER

We model the 7-state intraoperative physiological model and the PD controller as a hybrid system as illustrated in the Appendix, Figure 1. It is standard practice to perform perioperative monitoring of the patient to ensure the patient is stable enough for surgery. During the perioperative period (typically at least 30 minutes), if the patient exhibits extreme glucose variation, the surgery may be postponed until the patient stabilizes (Lipshutz and Gropper (2009)). To model the perioperative monitoring procedure, we divide the verification time into two phases: during the initial monitoring phase, if the glucose output y leaves a control range (e.g., $70-130$ mg/dL), the system transitions into the "NOT ADMIT" mode; if the glucose output y stays within the control range during the entire monitoring period, then the system transitions into the protocol control phase and the PD controller starts operating. During the protocol control phase, the system transitions into the "NOT SAFE" mode if the glucose output y leaves a safe range (e.g., $60 - 150$ mg/dL).

The hybrid system contains seven states: one initial state mode 0; four states (modes 1 - 4) that represents the system dynamics with four possible combinations of the two saturation switch terms in Equation 3a, which are restated in Equation 4; one "NOT ADMIT" mode and one "NOT SAFE" mode.

$$\begin{aligned} &\max(0, C_1), \ \textit{where } C_1 = k_{p1} - k_{p2} * G_p - k_{p3} * I_d \\ &\max(0, C_2), \ \textit{where } C_2 = k_{e1} * (G_p - k_{e2}) \end{aligned} \tag{4}$$

The system has 30 continuous states [4]

[3] Our previous paper Kohl et al. (2013) explains in detail the process of identifying the controller gains.

[4] To be consistent with the dReach implementation in Section 5, in the hybrid system model we denote all parameters as continuous states with derivatives of zero (i.e., constants).

Table 3. The PD Controller

Condition	Control Input Update
$y(k) \leq 60$	$uc(k) = 0, ub(k) = 0, m(k) = 12.5$
$60 < y(k) < 100$ AND $y(k) - y(k-1) < -30$	$uc(k) = 0, ub(k) = 0, m(k) = -0.1 * (y(k) - y(k-1))$
$100 \leq y(k) < 300$ OR $y(k) - y(k-1) \geq -30$	$uc(k) = \max(0, 0.05 * (y(k) - 100) + 0.06 * (y(k) - y(k-1))) + 1), ub(k) = 0, m(k) = 0$
$y(k) \geq 300$	$u(k) = 15, ub(k) = 15, m(k) = 0$

$$\mathcal{X} = \{I_p, X, I_1, I_d, I_l, G_p, G_t, \boldsymbol{P}, t, tau, y_{pre}, u, m\},$$

where $\boldsymbol{P}$ denotes the 18 model parameters, t is the global verification time, tau is the local timer variable, $y_{pre}(t)$ is a variable to record the last output sample, $u(t)$ and $m(t)$ are the insulin and meal inputs.

For simplicity of presentation we denote the four combinations of the two max terms using T_1 to T_4, as shown in Equation 5.

$$\begin{aligned} T_1 &:= (C_1 \leq 0) \wedge (C_2 \leq 0) \\ T_2 &:= (C_1 > 0) \wedge (C_2 \leq 0) \\ T_3 &:= (C_1 > 0) \wedge (C_2 > 0) \\ T_4 &:= (C_1 \leq 0) \wedge (C_2 > 0) \end{aligned} \tag{5}$$

Mode 0 is the initial state, in which all states have zero derivatives except t and tau. The system immediately goes into one of modes 1 - 4. The invariant on mode 0 is $INV_0 := (tau \leq 0)$. Equation 6 defines the guards on the transitions out of mode 0.

$$\forall i \in \{1, 2, 3, 4\},\ G[0 \rightarrow i] := T_i \wedge (tau \geq 0) \tag{6}$$

Let $t \in [0, t_a]$ denote the monitoring phase. Let $\mathcal{R}_{na}$ and $\mathcal{R}_{unsafe}$ denote the set of "NOT ADMIT" glucose values and "NOT SAFE" glucose values, respectively. Equation 7 defines the invariants on modes 1 - 4. To model the practical scenario that a clinician may not check exactly at the 30 minutes mark, we allow timing non-determinism by relaxing the conditions on the invariants with a sampling jitter δ.

$$\begin{aligned} \forall i \in \{1, 2, 3, 4\}, INV_i :=& (\neg(t \leq t_a \wedge y \in \mathcal{R}_{na}) \\ & \wedge (\neg(t > t_a \wedge y \in \mathcal{R}_{unsafe}) \\ & \wedge T_i \\ & \wedge (tau \leq 30 + \delta)) \end{aligned} \tag{7}$$

The self-transitions on modes 1 - 4 are triggered at the glucose sample times. On the self-transitions $\forall i \in \{1, 2, 3, 4\},\ G[i \rightarrow i]$, control inputs u and m are updated according to the PD algorithm, and y_{pre} is updated to the current y. Considering the timing jitter δ, Equation 8 defines the self-transition guards.

$$\forall i \in \{1, 2, 3, 4\},\ G[i \rightarrow i] := (tau \geq 30 - \delta) \tag{8}$$

The transition guards between modes 1 - 4 are governed by conditions T_1 - T_4 and are defined in Equation 9.

$$\forall i, j \in \{1, 2, 3, 4\},\ G[i \rightarrow j] := T_j \tag{9}$$

In modes 1 - 4, if $y \in \mathcal{R}_{na}$ during the monitoring phase, the system transitions into the "NOT ADMIT" mode 5. Equation 10 defines the transition guards between modes 1 - 4 and the "NOT ADMIT" mode 5.

$$\forall i \in \{1, 2, 3, 4\},\ G[i \rightarrow 5] := (t \leq t_a \wedge y \in \mathcal{R}_{na}) \tag{10}$$

In modes 1 - 4, if $y \in \mathcal{R}_{unsafe}$ after the monitoring phase, the system transitions into the "NOT SAFE" mode 6. Equation 11 defines the transition guards between modes 1 - 4 and the "NOT SAFE" mode 6.

$$\forall i \in \{1, 2, 3, 4\},\ G[i \rightarrow 6] := (t > t_a \wedge y \in \mathcal{R}_{unsafe}) \tag{11}$$

The "NOT ADMIT" mode 5 and "NOT SAFE" mode 6 are terminating states with no invariants or transitions out of them. The safety verification question is specified as follows: for all initial conditions (where the 7 physiological states and 18 parameters are in their ranges), can the system reach the "NOT SAFE" mode 6.

5. CASE STUDY: VERIFICATION OF A GLUCOSE CONTROL PROTOCOL

Verifying the intraoperative glucose controller safety property requires either a tool designed for non-linear dynamics, e.g., Flow* (Chen et al. (2013)), KeYmaera (Platzer and Quesel (2008)), and dReach/dReal (Kong et al. (2015)), or transforming the non-linear hybrid automata into a form suitable for other tools, e.g., UPPAAL (Larsen et al. (1997)), HyTech (Henzinger et al. (1997)), PHAVer (Frehse (2005)), and SpaceEx (Frehse et al. (2011)). While evaluating all the aforementioned verification tools against the intraoperative glucose control benchmark can provide useful insight to their respective capabilities, it is beyond the scope of this work. Rather, we provide a proof-of-concept illustration that for at least one verification tool, dReach, it is possible to verify the safety property over a non-scalar subspace of the potential patient physiology. The remainder of this section provides a brief description of the dReach implementation and summarizes the verification results.

5.1 Benchmark implementation in dReach

The dReach approach utilizes the framework of δ-complete decision procedures that aims to solve first-order logic formula with arbitrary computable real functions (Gao et al. (2013b)). The dReach tool can be employed to prove safety properties of hybrid systems over finite time by identifying safe and unsafe regions of the state space and defining a corresponding δ-decision problem. Following Gao et al. (2013b), we consider the δ-*decision* problem

$$\begin{aligned} \exists \mathcal{X}_{init} \wedge \exists t \in [0, t_{max}] \wedge \exists y \in \mathcal{R}_{unsafe} \text{s.t.} \\ |\mathcal{X}_{init}| \leq \delta_1 \wedge |y - \phi(t; \mathcal{X}_{init})| \leq \delta_2 \end{aligned} \tag{12}$$

where δ_i is a numerical error bound specified by an arbitrary rational number and the bounded first-order sentences contain Type 2 computable functions Ko (1991).

In this work, we define an unsafe region via limits on the glucose levels observable in the patient. We seek to show

that for our controller, composed with the physiological model described by a hybrid system with non-linear ODEs, there does not exist an initial condition which can lead to the satisfiability of (12) within a bounded time. As a conservative solution, the dReach tool (through δ-*weakening*) verifies for all initial conditions and bounded time that either the unsafe region is unreachable (UNSAT) or the unsafe region is reachable within a δ error (δ-SAT).

The dReach implementation of the surgical glucose hybrid system contains 30 state variables: 7 physiological; 18 parameters; 2 inputs (insulin rate u and glucose rate m); 1 state to record the last glucose reading; 1 global time state, and 1 local timer state. The dReach source code of this implementation is available online [5] .

5.2 Verification Results using dReach

To perform verification, we employ dReach version 2.15.01 on a Linux server with a Intel(R) Xeon(R) E5-2667 v2 3.30GHz CPU and 64 GB memory, and the results are provided in Table 4. First, we note that dReach is a bounded model checker, therefore the search depth or *Path Length* refers to the number of discrete transitions for which we have performed verification. In the results, the *Path Length* is the search depth completed by dReach in *Time* concluding in *Result*, where DNF translates to *did not finish* and $\boldsymbol{x}_0$ and $\boldsymbol{p}_0$ denote the nominal states and parameters specified in Table 1 and Table 2, respectively. From the results we observe that allowing the parameters and initial state to vary fully over their respective ranges prevents dReach from reaching a depth of more than 3. In this scenario, a path length of 3 corresponds to a maximum of one hour of surgery. The fact that dReach did not exceed the arguably trivial depth of 3 suggests that fully varying the parameter and initial condition space is a computationally challenging problem.

Table 4. Verification Results for $\mathcal{R}_{safe} = [60, 180]$.

Physiological Range		Path	Time	Result
State	Parameter	Length	(hours)	
Full	Full	3	30	safe
Full	Full	4	DNF	-
Full	$\boldsymbol{p}_0$	3	0.1	safe
Full	$\boldsymbol{p}_0$	4	0.6	safe
Full	$\boldsymbol{p}_0$	5	3.1	safe
Full	$\boldsymbol{p}_0$	6	8.2	safe
Full	$\boldsymbol{p}_0$	7	16.4	safe
Full	$\boldsymbol{p}_0$	8	DNF	-
$\boldsymbol{x}_0 \pm 0.5$	$\boldsymbol{p}_0 \pm 0.5$	3	0.1	safe
$\boldsymbol{x}_0 \pm 0.5$	$\boldsymbol{p}_0 \pm 0.5$	4	0.4	safe
$\boldsymbol{x}_0 \pm 0.5$	$\boldsymbol{p}_0 \pm 0.5$	5	1.1	safe
$\boldsymbol{x}_0 \pm 0.5$	$\boldsymbol{p}_0 \pm 0.5$	6	2.9	safe
$\boldsymbol{x}_0 \pm 0.5$	$\boldsymbol{p}_0 \pm 0.5$	7	8.1	safe
$\boldsymbol{x}_0 \pm 0.5$	$\boldsymbol{p}_0 \pm 0.5$	8	DNF	-

To investigate the capabilities of dReach, we allowed the initial state to vary over the full range, but constrained the parameters to equal $\boldsymbol{p}_0$. These results are consistent with the T1DMS scenario for a single artificial patient with unknown initial state (but known parameters). Here we observe a significant improvement in verification results, with dReach achieving a depth of 7 in 16.4 hours corresponding to a maximum surgery duration of 3.5 hours. By constraining the initial variance of the state and parameters to a hypercube around the nominal patient, we observe that dReach is able to achieve a depth of 7 in 8.1 hours corresponding to a maximum surgery duration of 3.5 hours. This suggests that discretizing the parameter and initial condition space can significantly improve time-to-verification given sufficient computing resources.

[5] URL:https://github.com/chen333/igc-benchmark

6. DISCUSSION AND FUTURE WORK

In this work, we consider the problem of safety verification for an intraoperative glucose controller. We present a formal model of the combined physiology and controller as a medical verification benchmark containing non-linear dynamics and over 30 states and parameters combined. Using dReach, a powerful non-linear verification tool, we provide preliminary results illustrating its performance on the proposed benchmark. Future work includes continued attempts to formally verify the proposed benchmark over the entire physiological space for a surgical duration of several hours (consistent with typical operations). Motivated by the stability inherent in biological systems, we plan to investigate methods to improve the verification performance through Lyapunov bounding of the state dynamics.

ACKNOWLEDGEMENTS

The authors would like to thank Soonho Kong at Carnegie Mellon University and Sicun Gao at Massachusetts Institute of Technology for helping us better understand and use the dReach tool for our verification problem. The authors would also like to thank Jim Kapinski at Toyota for sharing important experience and insights on using the dReach/dReal tool. This research is supported in part by NSF CNS-1035715 and in part by the DGIST Research and Development Program of the Ministry of Science, ICT and Future Planning of Korea (CPS Global Center).

REFERENCES

Basu, R., Di Camillo, B., Toffolo, G., Basu, A., Shah, P., Vella, A., Rizza, R., and Cobelli, C. (2003). Use of a novel triple-tracer approach to assess postprandial glucose metabolism. *American Journal of Physiology-Endocrinology And Metabolism*, 284(1), E55–E69.

Bochicchio, G.V., Sung, J., Joshi, M., Bochicchio, K., Johnson, S.B., Meyer, W., and Scalea, T.M. (2005). Persistent hyperglycemia is predictive of outcome in critically ill trauma patients. *Journal of Trauma and Acute Care Surgery*, 58(5), 921–924.

Bruno, A., Gregori, D., Caropreso, A., Lazzarato, F., Petrinco, M., and Pagano, E. (2008). Normal glucose values are associated with a lower risk of mortality in hospitalized patients. *Diabetes Care*, 31(11), 2209–2210.

Chen, S., Weimer, J., Rickels, M., Peleckis, A., and Lee, I. (2015). Towards a model-based meal detector for type i diabetics. In *the Medical Cyber Physical Systems Workshop, CPS Week, Seattle, WA*.

Chen, X., Ábrahám, E., and Sankaranarayanan, S. (2013). Flow*: An analyzer for non-linear hybrid systems. In *Computer Aided Verification*, 258–263. Springer.

Cobelli, C., Renard, E., and Kovatchev, B. (2011). Artificial pancreas: past, present, future. *Diabetes*, 60(11), 2672–2682.

Dalla Man, C., Micheletto, F., Lv, D., Breton, M., Kovatchev, B., and Cobelli, C. (2014). The uva/padova type 1 diabetes simulator new features. *Journal of diabetes science and technology*, 8(1), 26–34.

Danaei, G., Finucane, M.M., Lu, Y., Singh, G.M., Cowan, M.J., Paciorek, C.J., Lin, J.K., Farzadfar, F., Khang, Y.H., Stevens, G.A., et al. (2011). National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants. *The Lancet*, 378(9785), 31–40.

Frehse, G. (2005). Phaver: Algorithmic verification of hybrid systems past hytech. In *Hybrid Systems: Computation and Control*, 258–273. Springer.

Frehse, G., Le Guernic, C., Donzé, A., Cotton, S., Ray, R., Lebeltel, O., Ripado, R., Girard, A., Dang, T., and Maler, O. (2011). Spaceex: Scalable verification of hybrid systems. In *Computer Aided Verification*, 379–395. Springer.

Gao, S., Kong, S., and Clarke, E.M. (2013a). dreal: An smt solver for nonlinear theories over the reals. In *Automated Deduction–CADE-24*, 208–214. Springer.

Gao, S., Kong, S., and Clarke, E.M. (2013b). Satisfiability modulo odes. In *Formal Methods in Computer-Aided Design (FMCAD), 2013*, 105–112. IEEE.

Harris, M.I., Hadden, W.C., Knowler, W.C., and Bennett, P.H. (1987). Prevalence of diabetes and impaired glucose tolerance and plasma glucose levels in us population aged 20–74 yr. *Diabetes*, 36(4), 523–534.

Henzinger, T.A., Ho, P.H., and Wong-Toi, H. (1997). Hytech: A model checker for hybrid systems. In *Computer aided verification*, 460–463. Springer.

Katz, A., Nambi, S.S., Mather, K., Baron, A.D., Follmann, D.A., Sullivan, G., and Quon, M.J. (2000). Quantitative insulin sensitivity check index: a simple, accurate method for assessing insulin sensitivity in humans. *The Journal of Clinical Endocrinology & Metabolism*, 85(7), 2402–2410.

Ko, K.I. (1991). *Complexity theory of real functions*. Birkhauser Boston Inc.

Kohl, B.A., Chen, S., Mullen-Fortino, M., and Lee, I. (2013). Evaluation and enhancement of an intraoperative insulin infusion protocol via in-silico simulation. In *Healthcare Informatics (ICHI), 2013 IEEE International Conference on*, 307–316. IEEE.

Kong, S., Gao, S., Chen, W., and Clarke, E.M. (2015). dreach: Delta-reachability analysis for hybrid systems. In *Tools and Algorithms for the Construction and Analysis of Systems - 21st International Conference, TACAS 2015, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2015, London, UK, April 11-18, 2015. Proceedings*, 200–205.

Kovatchev, B.P., Breton, M., Man, C.D., and Cobelli, C. (2009). In silico preclinical trials: A proof of concept in closed-loop control of type 1 diabetes. *Diabetes Sci Technol*, 3(1), 44–55.

Kulcu, E., Tamada, J.A., Reach, G., Potts, R.O., and Lesho, M.J. (2003). Physiological differences between interstitial glucose and blood glucose measured in human subjects. *Diabetes care*, 26(8), 2405–2409.

Laakso, M. (1993). How good a marker is insulin level for insulin resistance? *American Journal of Epidemiology*, 137(9), 959–965.

Larsen, K.G., Pettersson, P., and Yi, W. (1997). Uppaal in a nutshell. *International Journal on Software Tools for Technology Transfer (STTT)*, 1(1), 134–152.

Lazar, H.L., Chipkin, S.R., Fitzgerald, C.A., Bao, Y., Cabral, H., and Apstein, C.S. (2004). Tight glycemic control in diabetic coronary artery bypass graft patients improves perioperative outcomes and decreases recurrent ischemic events. *Circulation*, 109(12), 1497–1502.

Lipshutz, A. and Gropper, M.A. (2009). Perioperative glycemic control. *Anesthesiology*, 110(2), 408–21.

Man, C.D., Rizza, R.A., and Cobelli, C. (2007). Meal simulation model of the glucose-insulin system. *Biomedical Engineering and IEEE Transactions on*, 54(10), 1740–1749. doi: 10.1109/TBME.2007.893506.

McAlister, F.A., Majumdar, S.R., Blitz, S., Rowe, B.H., Romney, J., and Marrie, T.J. (2005). The relation between hyperglycemia and outcomes in 2,471 patients admitted to the hospital with community-acquired pneumonia. *Diabetes Care*, 28(4), 810–815.

Meijering, S., Corstjens, A.M., Tulleken, J.E., Meertens, J.H., Zijlstra, J.G., and Ligtenberg, J.J. (2006). Towards a feasible algorithm for tight glycaemic control in critically ill patients: a systematic review of the literature. *Critical Care*, 10(1), R19.

Platzer, A. and Quesel, J.D. (2008). Keymaera: A hybrid theorem prover for hybrid systems (system description). In *Automated Reasoning*, 171–178. Springer.

Subramaniam, B., Panzica, P., Novack, V., Mahmood, F., Matyal, R., Mitchell, J., Sundar, E., Bose, R., Pomposelli, F., Kersten, J., et al. (2009). Continuous perioperative insulin infusion decreases major cardiovascular events in patients undergoing vascular surgery: a prospective, randomized trial. *Anesthesiology*, 110(5), 970.

Wallymahmed, M., Dawes, S., Clarke, G., Saunders, S., Younis, N., and MacFarlane, I. (2005). Hospital in-patients with diabetes: increasing prevalence and management problems. *Diabetic Medicine*, 22(1), 107–109.

Xu, J., Kochanek, K.D., Murphy, S.L., and Tejada-Vera, B. (2010). Deaths: final data for 2007. *National vital statistics reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System*, 58(19), 1–19.

APPENDIX

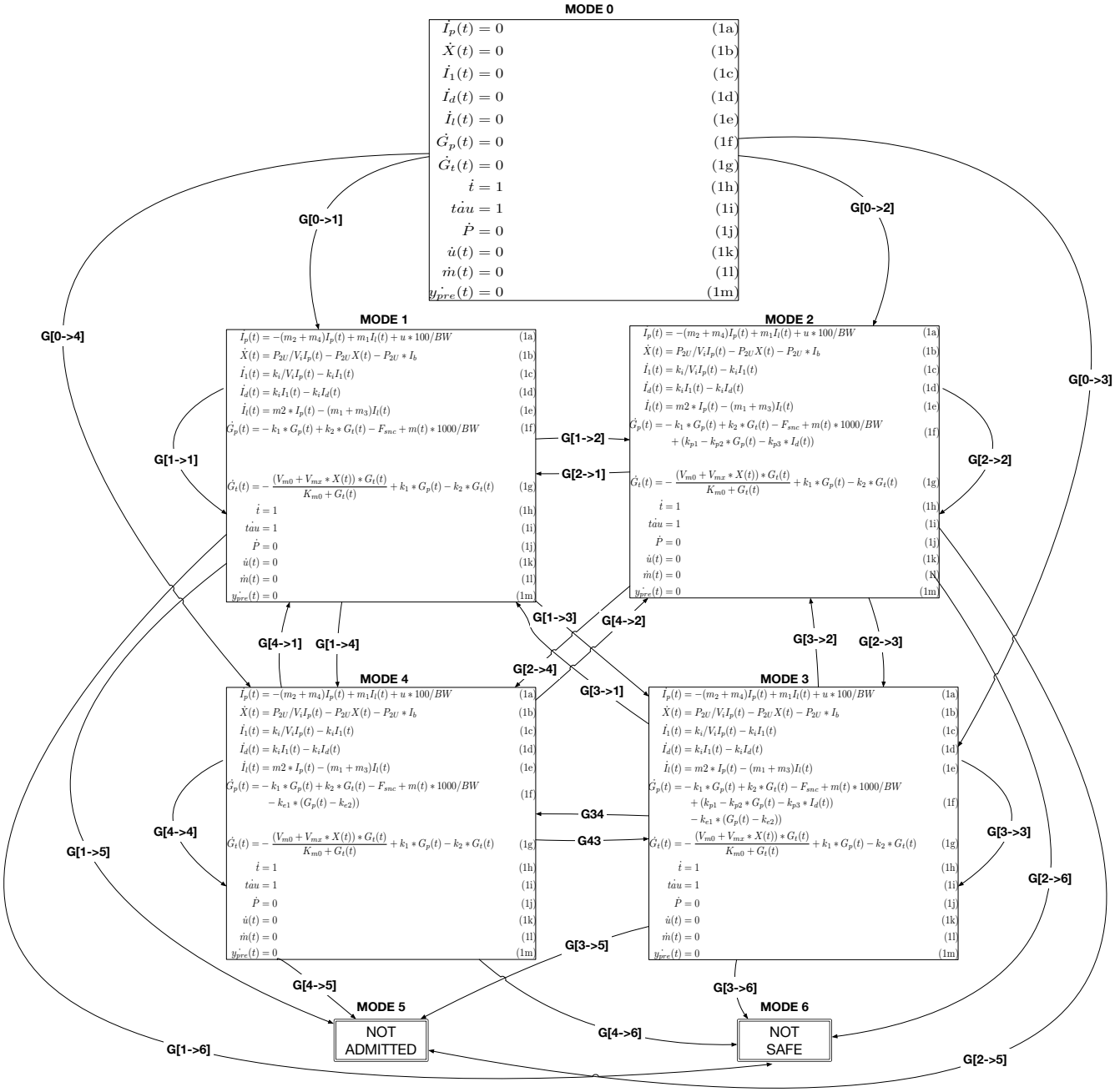

Fig. 1. A Hybrid System Representation of the FDA-accepted High-Fidelity Physiological Model with the PD Controller.